**COVID-19 cases analysis using IBM Cognos for visualization**.

Here are the steps you can follow to define your analysis objectives and obtain, process, and clean the COVID-19 data:

1. Define Analysis Objectives:

Before you start working with the data, it's essential to define your analysis objectives. What specific insights or information are you looking to gain from the COVID-19 data? Some common analysis objectives include:

- Tracking the spread of COVID-19 over time.

- Examining the relationship between COVID-19 cases and deaths.

- Identifying hotspots or regions with high infection rates.

- Visualizing vaccination progress alongside cases and deaths.

Clearly defining your objectives will guide your analysis and visualization efforts.

2. \*\*Obtain the COVID-19 Data:\*\*

You mentioned a dataset link from Kaggle. To obtain the data, follow these steps:

- Register for a Kaggle account if you don't already have one.

- Visit the provided dataset link (https://www.kaggle.com/datasets/chakradharmattapalli/covid-19-cases).

- Download the dataset from Kaggle to your local machine. It's usually available in a compressed format like a ZIP file or CSV.

3. \*\*Process and Clean the Data:\*\*

Data cleaning is a critical step to ensure the accuracy and reliability of your analysis. Here's how you can do it:

- \*\*Data Format\*\*: Check the format of the data you've downloaded (e.g., CSV). Ensure it's compatible with IBM Cognos, which can handle various data formats.

- \*\*Missing Data\*\*: Examine the dataset for missing or incomplete data. Decide how to handle missing values (e.g., impute or remove them).

- \*\*Data Types\*\*: Check the data types of each column and convert them if needed (e.g., dates to date format, numbers to appropriate numeric types).

- \*\*Duplicates\*\*: Remove any duplicate entries in the dataset.

- \*\*Inconsistent Data\*\*: Check for data inconsistencies or outliers that may skew your analysis. Decide whether to remove or correct them.

- \*\*Data Aggregation\*\*: If your dataset has daily data, consider aggregating it to weekly or monthly intervals, depending on your analysis objectives.

4. \*\*Load Data into IBM Cognos:\*\*

After cleaning and preparing the data, you can load it into IBM Cognos. IBM Cognos typically allows you to import data from various sources, including local files, databases, or cloud storage. Follow the steps in IBM Cognos to load your cleaned COVID-19 dataset.

5. \*\*Create Visualizations:\*\*

With the data loaded, you can now start building visualizations in IBM Cognos. Depending on your analysis objectives, you can create various types of charts and dashboards to present the data effectively. Common COVID-19 visualizations include line charts, bar charts, heatmaps, and geospatial maps.

6. \*\*Iterate and Refine:\*\*

The data analysis process often involves iterative steps. As you create visualizations, you might discover new insights or areas that require further cleaning or refinement. Be prepared to iterate and refine your analysis as needed.

Remember to document your analysis process and the steps you take to ensure transparency and reproducibility. IBM Cognos provides a range of tools for data analysis and visualization, making it a powerful platform for this type of project.

**VISUALIZATION:**

Data Visualization:

Use Python's data visualization libraries, such as Matplotlib or Seaborn, to create various types of visualizations. Here are some common visualization types:

**Line Charts:** Plot daily COVID-19 cases or deaths over time.

**Bar Charts**: Show the total cases or deaths for different countries or regions.

**Heatmaps:** Display COVID-19 data on a geographical map with color-coding for cases or deaths.

**Stacked Bar Charts**: Compare cases and deaths in different regions.

**Pie Charts:** Show the distribution of cases or deaths by region.

**Customize Visualizations:**

Customize your visualizations by adding titles, labels, legends, and other annotations to make them informative and visually appealing.







